Journal of Biomedical Informatics 44 (2011) S3–S4

Contents lists available at SciVerse ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Editorial

# Selected Papers from the 2011 Summit on Translational Bioinformatics

AMIA's Fourth Summit on Translational Bioinformatics (TBI Summit) was held in San Francisco from March 7–9, 2011. Dating back to 2008, the Summit has become recognized as a key venue for the exchange of ideas in translational bioinformatics. As in previous years, the 2011 TBI Summit highlighted the trans-disciplinary nature of this now-burgeoning research field and provided a unique opportunity to bring together the finest minds in translational bioinformatics from academia, industry, government, and non-profit sectors. The 2011 TBI Summit focused on topics aligned with four chosen themes: (1) informatics concepts, tools and techniques to enable integrative translational bioinformatics research; (2) informatics methods for the integrative analysis of molecular and clinical measurements; (3) relating and representing phenotypes and disease for translational bioinformatics research; and, (4) informatics methods bridging basic science discoveries and clinical practice.

In addition to the rich offering of scientific presentations (including 3 tutorials, 8 panels, 48 podium presentations, and 69 poster presentations), the 2011 TBI Summit featured a number of keynote presentations. Lincoln Stein, platform leader of Informatics and Bio-computing at the Ontario Institute for Cancer Research, delivered the opening keynote presentation on developing unified approaches for managing biological and clinical data associated with cancer. On the second day, award winning writer and lecturer Carl Zimmer gave a presentation on the Human Microbiome. Russ Altman, chair of bioengineering and professor at Stanford University, as well as the Principal Investigator for major TBI initiatives, such as the Pharmacogenomics Knowledge Base (PharmGKB), summarized the key advances in translational bioinformatics since the last (2010) TBI Summit (also summarized in the *Journal of the American Medical Informatics Association* [1]). Kenneth Buetow, director of the National Cancer Institute's Center for Bioinformatics and Information Technology, delivered the 'bridge day' keynote on developing national infrastructure to enable translational research (which also served as the opening keynote for the partner Summit on Clinical Research Informatics).

For the accepted full-length podium presentations, the top nine were chosen for publication based on ranking submissions according to reviewer submitted scores. The very best paper, as determined by an award committee (Russ Altman, Atul Butte, Riccardo Bellazzi, Andrea Califano, Isaac Kohane, Yves Lussier, and Peter Tarczy-Hornoch), received the first Marco F. Ramoni Award. The award was established to honor the memory of Marco Ramoni, who passed away suddenly in June 2010. Dr. Ramoni was a friend and colleague for many of us in the translational bioinformatics field as well as a major force behind the TBI Summit since its inception. An extended version of this prize-winning paper was published in the *Journal of the American Medical Informatics Association* [2]. The 2011 TBI Summit Track Chairs (Maricel Kann, Lewis Frey, Nigam Shah, and Jessica Tenenbaum) vetted the remaining eight distinguished paper candidates and they were subjected to another round of full peer-review and revision before being accepted for inclusion in this supplement. These eight papers describe key informatics innovations that aim to directly address challenges faced by the translational bioinformatics community.

The identification of genes of interest or biomarkers associated with complex disorders, such as autism spectrum disorders or breast cancer, is a perennial challenge in the biomedical research community. Meehan et al. [3] describe an ontology-based method for determining candidate autism genes in the popular model organism, *Mus musculus* (house mouse). Fabregue et al. [4] then describe an approach to identify molecular biomarkers that might be used as prognostic or diagnostic tools in the context of breast cancer. These studies demonstrate how bioinformatics approaches may offer promise towards a better understanding of the molecular underpinnings of disease, while still highlighting the key challenges that remain.

The identification of key genes or features embodies a major area in translational bioinformatics. However, the complexity of many diseases or disorders requires robust validation as well as new approaches for data exploration. Ganchev et al. [5] present a framework for discovering biomarkers from multiple data sets simultaneously. Bhavnani et al. [6] showcase network analysis approaches to explore co-occurrence relationships between cytokines in asthma patients. As data acquisition techniques continue to produce volumes of new data, such approaches for knowledge integration and data exploration may become essential tools for the translational bioinformatician.

Representation of biological and disease knowledge is an essential component in many translational bioinformatics studies. LePendu et al. [7] describe a feasibility study to develop a gene-disease annotation data set (akin to the publicly available Gene Ontology Annotation data set) for enabling statistical ranking of disease genes. The promising findings by Ruau et al. [8] suggest that automated tools may offer some solace in the data deluge of largely un-annotated data sets. Scotch et al. [9] then demonstrate how existing stores of biological knowledge (e.g., GenBank) can be leveraged for secondary studies (e.g., phylogeography). Finally, Tran et al. [10] describe a system that leverages ontological approaches for representing disease information from multiple clinical trial sites.

Collectively, the papers showcased here reflect the great advances that translational bioinformatics is making towards gaining a deeper understanding of the mechanisms associated with human disease. The major thrust still remains in the development of approaches to manage the large volumes of data that are being


doi:10.1016/j.jbi.2011.11.014

generated in the current era of higher and higher throughput technologies. However, the findings of all the papers presented here echo a major theme that resounded across all the presentations at the 2011 Summit on Translational Bioinformatics: *Translational Bioinformatics is increasingly focusing the rich array of bioinformatics approaches towards a better understanding of human disease that may ultimately lead to a new generation of prophylaxes, diagnostics, and therapeutics.*

## References

[1] Altman RB, Miller KS. 2010 translational bioinformatics year in review. J Am Med Inform Assoc 2011;18(4):358–66.
[2] Wei W, Visweswaran S, Cooper GF. The application of naive Bayes model averaging to predict Alzheimer's disease from genome-wide data. J Am Med Inform Assoc 2011;18(4):370–5.
[3] Meehan TF, Carr CJ, Jay JJ, Bult CH, Chesler EJ, Blake JA. Autism candidate genes via mouse phenomics. J Biomed Info 2011;44(S1):S5–S11.
[4] Fabregue M, Bringay S, Poncelet P, Teisseire M, Orsetti B. Mining microarray data to predict the histological grade of a breast cancer. J Biomed Info 2011;44(S1):S12–6.
[5] Ganchev P, Malehorn D, Bigbee WL, Gopalkrishnan V. Transfer learning of classification rules for biomarker discovery and verification from molecular profiling studies. J Biomed Info 2011;44(S1):S17–23.
[6] Bhavnani SK, Victor S, Calhoun WJ, Busse WW, Bleeker E, Castro, et al. How cytokines co-occur across asthma patients: From bipartite network analysis to a molecular-based classification. J Biomed Info 2011;44(S1):S24–30.
[7] LePendu P, Musen MA, Shah NH. Enabling enrichment analysis with the human disease ontology. J Biomed Info 2011;44(S1):S31–8.
[8] Ruau D, Mbagwu M, Dudley JT, Krishnan V, Butte AJ. Comparison of automated and human assignment of MeSH terms on publicly-available molecular datasets. J Biomed Info 2011;44(S1):S39–43.
[9] Scotch M, Sarkar IN, Mei C, Leaman R, Cheung KH, Ortiz P, et al. Enhancing phylogeography by improving geographical information from GenBank. J Biomed Info 2011;44(S1):S44–7.
[10] Tran VA, Johnson N, Redline S, Zhang GQ. OnWARD: ontology-driven web-based framework for multi-center clinical studies. J Biomed Info 2011;44(S1):S48–53.

Indra Neil Sarkar
*Center for Clinical and Translational Science, University of Vermont, Burlington, VT, USA*

*Department of Microbiology and Molecular Genetics, University of Vermont, Burlington, VT, USA*

*Department of Computer Science, University of Vermont, Burlington, VT, USA*